

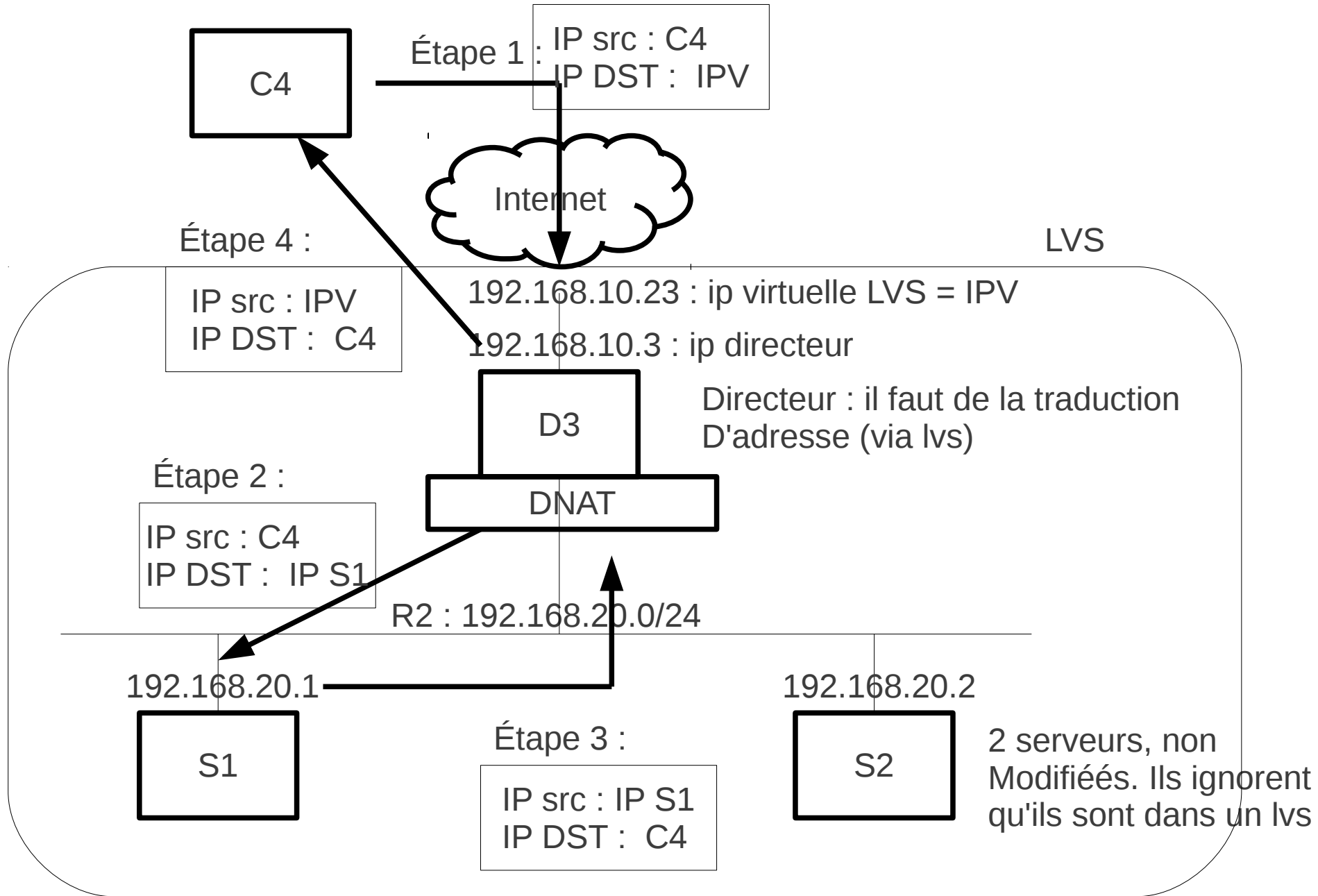
Équilibrage de charge

- Linux virtual server
 - Un directeur répartit les requêtes
 - Les serveurs ne le savent pas forcément
 - Agit comme un répartiteur de couche 4
- Cluster beowulf
 - La charge est répartie de façon coopérative entre les serveurs du cluster

Principe de fonctionnement

- Les clients voient l'IP virtuelle (VIP) du LVS
- La VIP est routable et accessible depuis les clients
- Les serveurs ont des IP réelles (RIP)
- Le directeur est un répartiteur de charge (load balancer)
- 3 modes de fonctionnement :
 - LVS-NAT
 - LVS-DR
 - LVS-TUN

LVS-NAT



LVS-NAT

- 1) Le client fait une requête sur l'ip virtuelle du LVS (IPV) ;
- 2)DNAT : Le directeur change l'ip destination (et parfois le port destination) qui devient celle du serveur cible (IP S1 dans l'exemple) ;
- 3)Le serveur répond au client. Sa réponse transite par le directeur qui est aussi son routeur par défaut ;
- 4)La règle DNAT de l'étape2 est appliquée dans l'autre sens et le paquet est transmis au client avec ipv comme ip source.

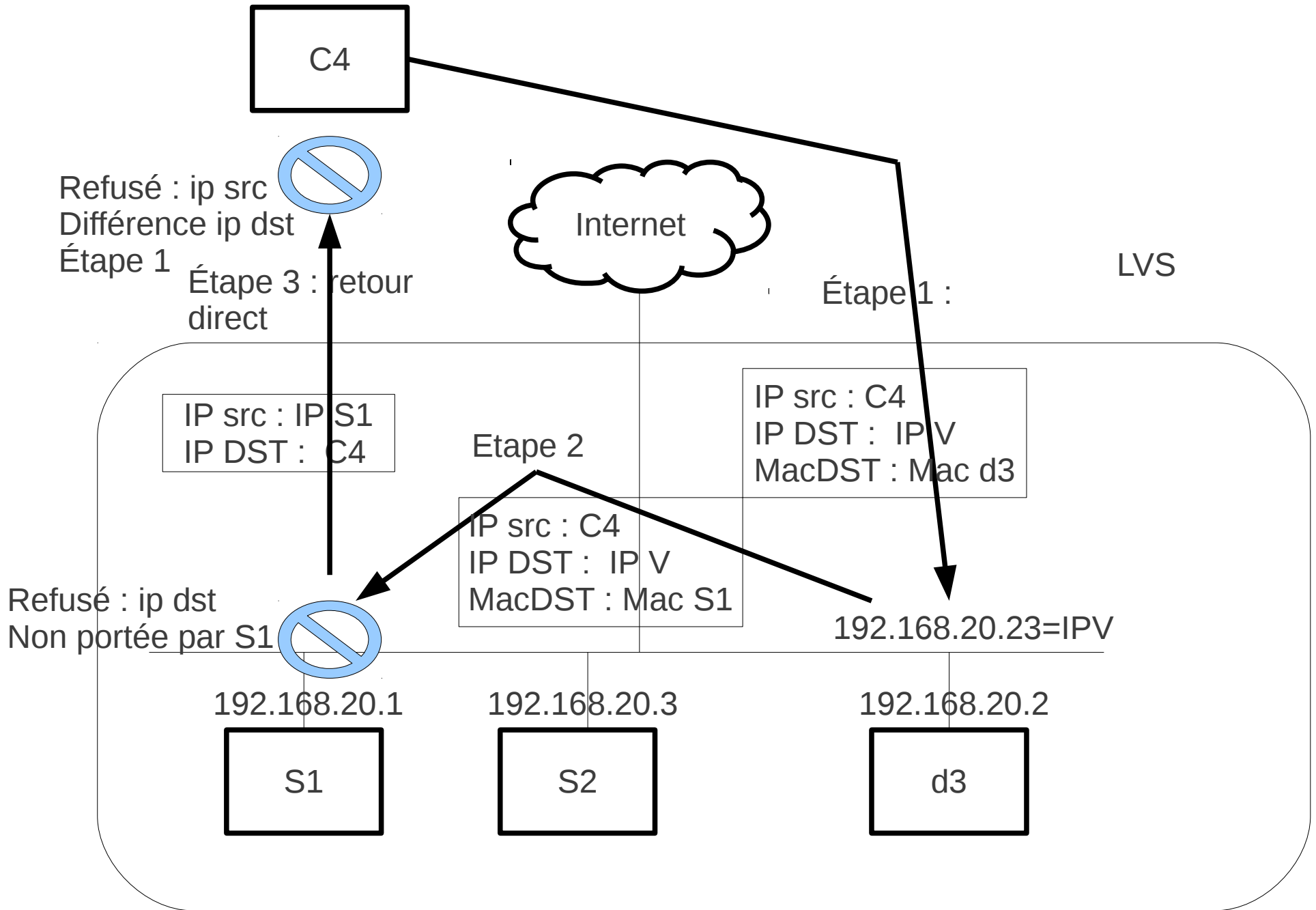
LVS-NAT

- Les clients interagissent avec le directeur
- Le directeur fait de la traduction d'adresse
- Le directeur est le routeur de sortie des serveurs
- Le DNAT est géré par le module ip_vs
- Avantage :
 - les serveurs ignorent qu'ils sont dans un LVS
 - OS des serveurs sans importance
 - Les serveurs WeB peuvent écouter sur un port différent de celui sur lequel le client envoie sa requête (permet par le DNAT)
- Défauts :
 - Tout passe par le directeur en entrée (petit) et en sortie (gros)

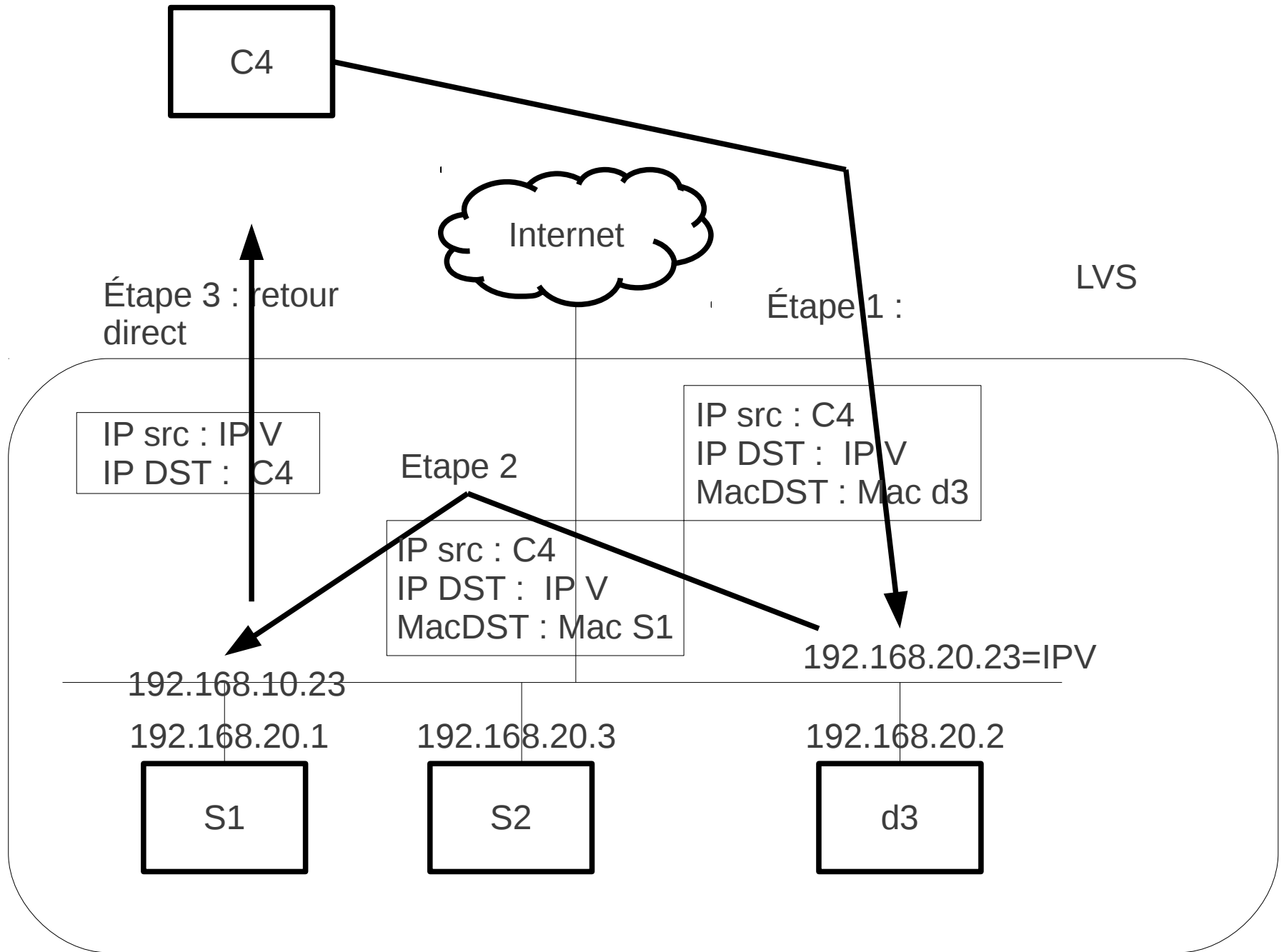
LVS-DR

- But : éviter de passer par le directeur pour les réponses des serveurs
- Le retour se fait directement depuis les serveurs vers les clients
- Depuis une ip virtuelle commune aux serveurs et au directeur
- Seul le directeur annonce cette ip via arp (car lui seul doit avoir les paquets en entrée)

LVS-DR



LVS-DR : solution



LVS-DR : arp sur les serveurs

- Seul le directeur doit répondre aux requêtes ARP sur l'ip virtuelle du LVS
- Solutions :
 - Affecter l'IPV à chaque serveur et :
 - Paramètres arp_announce et arp_ignore du noyau des server
 - Filtrage par arptables :
 - Filtrer sur les serveur tout paquet ARP entrant destiné à l'ip virtuelle
 - Remplacer l'ip source virtuelle des annonces arp des serveurs par leur ip réelle
 - En faisant de la redirection d'adresse (éventuellement de port) en entrée sur les serveurs avec netfilter/iptables
 - Tout ce qui est destiné à l'ip virtuelle est réécrit comme destiné à l'ip réelle du serveur
 - Dans ce cas, le serveur n'a pas besoin d'avoir l'ip virtuelle comme ip

LVS-TUN

- Idem lvs-DR mais le lien directeur-serveurs est un tunnel
- N'impose pas aux serveurs d'être localisés au même endroit

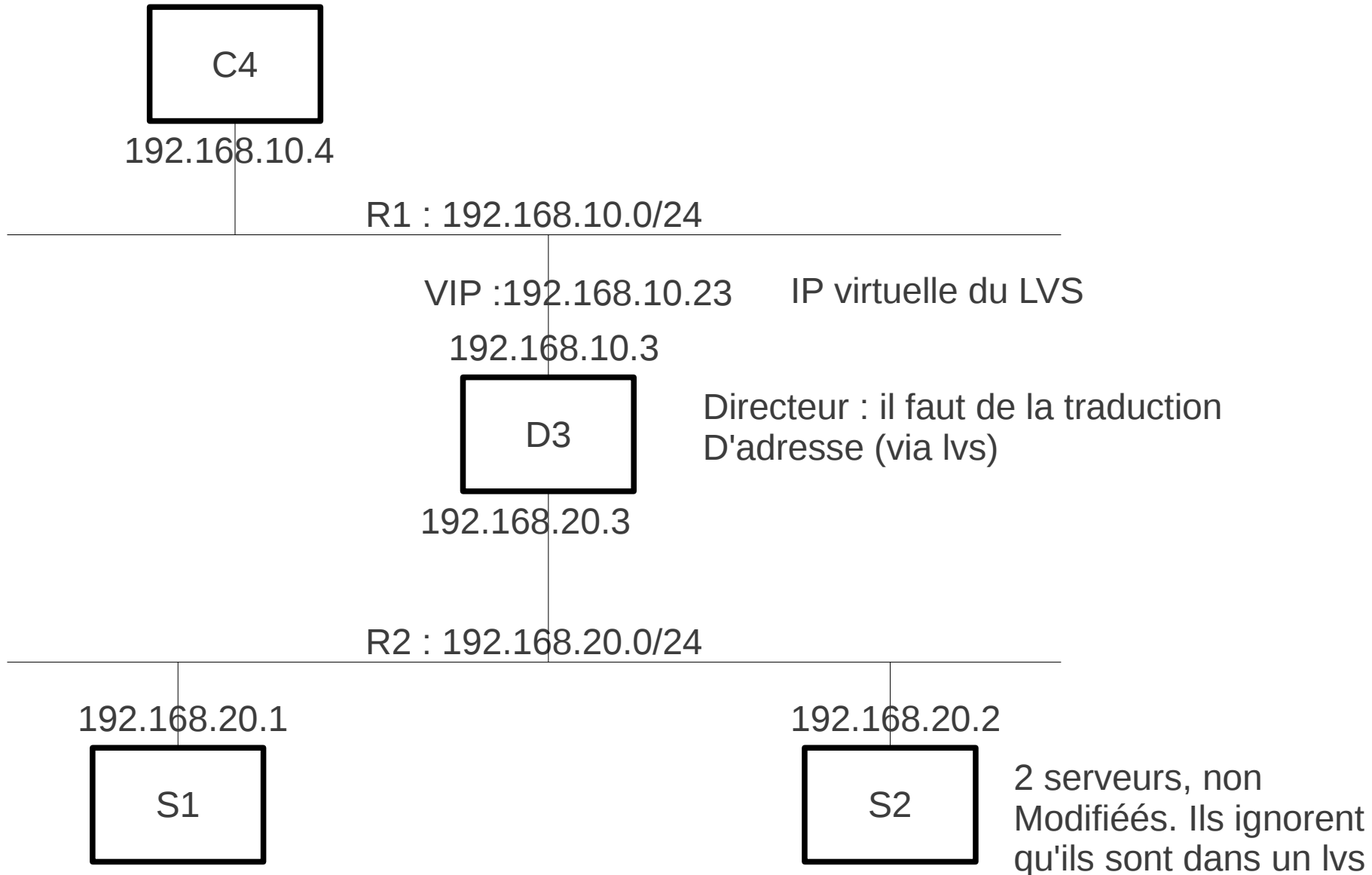
Haute disponibilité/répartition de charge

- LVS : palie l'indisponibilité d'un serveur
- LVS : répartit la charge sur les serveur
- MAIS :
 - Le directeur est un SPOF
 - Le directeur assume toute la charge de son travail de directeur
- Solution :
 - Haute disponibilité sur le directeur (directeur backup, ...)

Pré-requis logiciels

- Noyau supportant lvs
 - Dans le noyau : `grep ip_vs System.map`
 - En module :
 - Chargé à la première utilisation de `ip_vs`
 - La disponibilité peut-être vérifié avec `modinfo ip_vs`
- Commande `ipvsadm` (package debian `ipvsadm`)

Maquette de test LVS-NAT



Configuration de base

- Sur S1 : un serveur WeB écoutant sur le port 80 avec une page d'accueil contenant le texte « Je suis le serveur 1 et j'ai faim »
- Sur S2 : un serveur WeB écoutant sur le port 80 avec une page d'accueil contenant le texte « Je suis le serveur 2 et j'ai soif»
- Le directeur doit être le routeur par défaut des deux serveurs
- Activation du routage sur le directeur :
 - Décommenter la ligne `net.ipv4.ip_forward` dans `/etc/sysctl.conf`
 - Puis recharger le fichier par « `sysctl -p` »

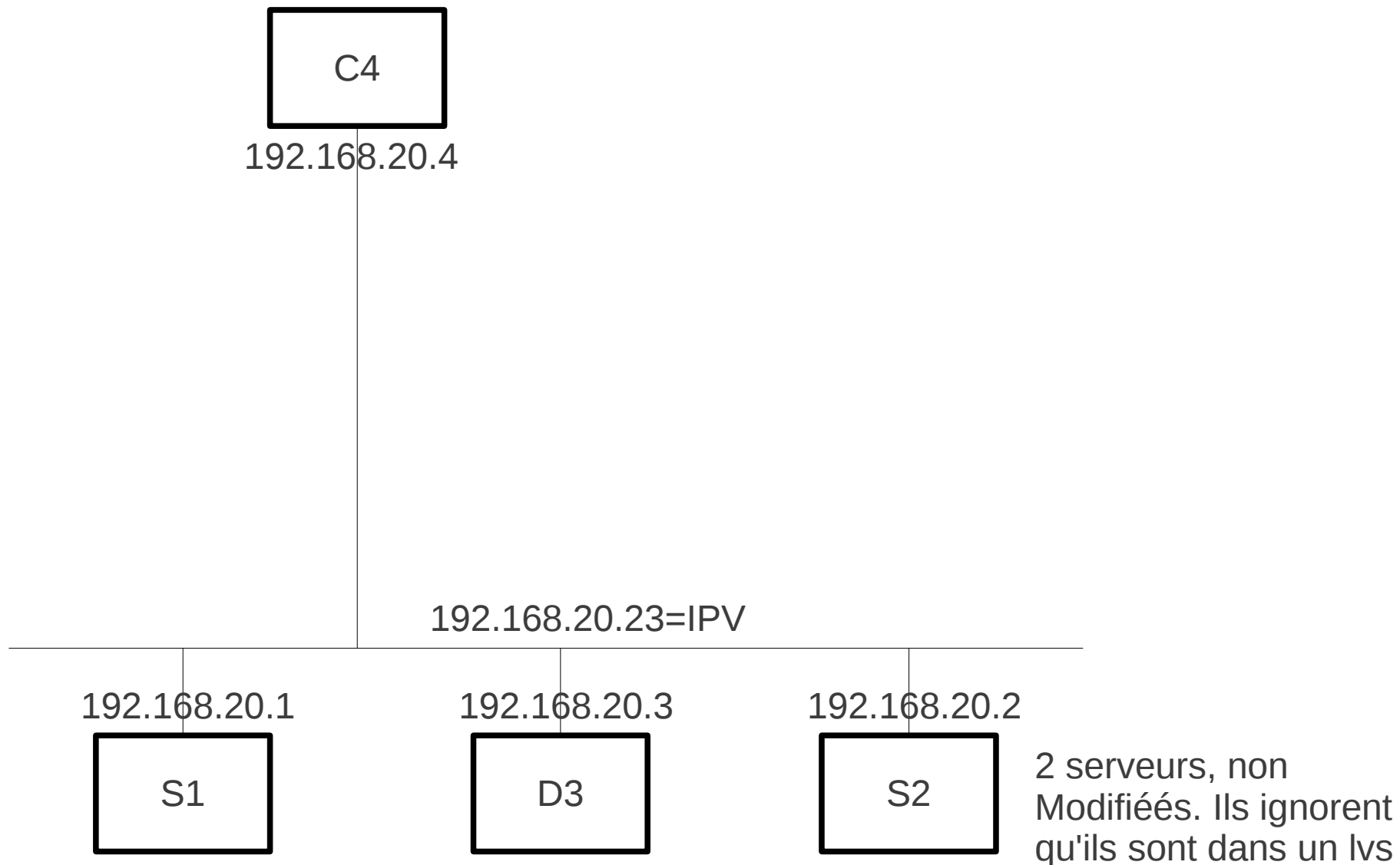
Configuration de LVS-NAT

- Déclaration du LVS :
 - `ipvsadm -A -t 192.168.10.23:80 -s rr`
 - `-A` : ajout d'un LVS
 - `-t` : tcp
 - `-s` : méthode de répartition de charge
- Ajout des serveurs
 - `ipvsadm -a -t 192.168.10.23:80 -r 192.168.20.1:80 -m -w 1`
 - `-a` : ajout d'un serveur réel à un LVS
 - `-r` : ip:port du serveur réel à ajouter
 - `-m` : active la traduction d'adresse (masquerading)
 - `-w` : poids (inutile ici car on est en round-robin)

Test de LVS-NAT

- Avec un navigateur, faites des requêtes sur <http://192.168.10.23/> (ip virtuelle du LVS)
- Commentez l'affichage
- Sur le LVS, que donne la commande `ipvsadm ? ipvsadm -L ? ipvsadm -Lc ?`
- Faites une requête et consultez encore `ipvsadm` (`ipvsadm -L, -Lc, -Ln, ...`)
- Faites des analyses de trames sur les deux réseaux et comparez les trames avant et après
- Déconnectez la carte réseau vmware d'un des serveurs et répétez les étapes ci-dessus

Maquette de test LVS-DR



Configuration de base

- Sur S1 : un serveur WeB écoutant sur le port 80 avec une page d'accueil contenant le texte « Je suis le serveur 1 et j'ai faim »
- Sur S2 : un serveur WeB écoutant sur le port 80 avec une page d'accueil contenant le texte « Je suis le serveur 2 et j'ai soif»
- Le directeur a l'IP virtuelle du LVS comme seconde IP sur sa carte réseau
- Vous appliquerez l'une des trois solutions (ou les trois tour à tour en fonction de votre temps) pour que S1 et S2 acceptent les paquets destinés à IPV sans émettre de requêtes ARP indiquant l'IP IPV.

Configuration de LVS-DR

- Déclaration du LVS :
 - `ipvsadm -A -t 192.168.20.23:80 -s rr`
 - `-A` : ajout d'un LVS
 - `-t` : tcp
 - `-s` : méthode de répartition de charge
- Ajout des serveurs
 - `ipvsadm -a -t 192.168.20.23:80 -r 192.168.20.1:80 -g -w 1`
 - `-a` : ajout d'un serveur réel à un LVS
 - `-r` : ip:port du serveur réel à ajouter
 - `-g` : active le 'direct routing'
 - `-w` : poids (inutile ici car on est en round-robin)

Test de LVS-DR

- Avec un navigateur, faites des requêtes sur `http://192.168.20.23/` (ip virtuelle du LVS)
- Commentez l'affichage
- Sur le LVS, que donne la commande `ipvsadm` ?
`ipvsadm -L` ? `ipvsadm -Lc` ?
- Faites une requête et consultez encore `ipvsadm`
(`ipvsadm -L`, `-Lc`, `-Ln`, ...)
- Faites des analyses de trames sur les deux réseaux et comparez les trames avant et après
- Déconnectez la carte réseau vmware d'un des serveurs et répétez les étapes ci-dessus